Using that latter technology, DNA is carried through the cell wall and into the cytoplasm on the surface of small metal particles as described (Klein *et al.*, 1987; Klein *et al.*, 1988; McCabe *et al.*, 1988). The metal particles penetrate through several layers of cells and thus allow the transformation of cells within tissue explants.

5

4.10.4 GENE EXPRESSION IN PLANTS

The fact that plant codon usage more closely resembles that of humans and other higher organisms than unicellular organisms, such as bacteria, unmodified bacterial genes are often poorly expressed in transgenic plant cells. The apparent overall preference for

10   GC content in codon position three has been described in detail by Murray *et al.* (1990). The 207 plant genes described in this work permitted the compilation of codon preferences for amino acids in plants. These authors describe the difference between codon usage in monocots and dicots, as well as differences between chloroplast encoded genes and those which are nuclear encoded. Utilizing the codon frequency tables

15   provided, those of skill in the art can engineer such a bacterial sequence for expression in plants by modifying the DNA sequences to provide a codon bias for G or C in the third position. The reference provides an exhaustive list of tables to guide molecular geneticists in preparing synthetic gene sequences which encode the polypeptides of the invention, and which are expressed in transformed plant cells in a suitable fashion to

20   permit synthesis of the polypeptide of interest *in planta.*

A similar work by Diehn *et al.* (1996) details the modification of prokaryotic-derived gene sequences necessary to permit expression in plants.

Iannacone *et al.* (1997) describe the transformation of egg plant with a genetically engineered *B. thuringiensis* gene encoding a *cry*3 class endotoxin. Utilizing sequences

25   which avoid polyadenylation sequences, ATTA sequences, and splicing sites a synthetic gene was constructed which permitted expression of the encoded toxin *in planta.*

Expression of heterologous proteins in transgenic tobacco has been described by Rouwendal *et al.* (1997). Using a synthetic gene, the third position codon bias for C+G was created to permit expression of the jellyfish green fluorescent protein-encoding gene

30   *in planta.*

-88-

Fütterer and Hohn (1996) describe the effects of mRNA sequence, leader sequences, polycistronic messages, and internal ribosome binding site motis, on expression in plants. Modification of such sequences by construction of synthetic genes permitted expression of viral mRNAs in transgenic plant cells.

5   Preparation of transgenic plants which express genes encoding non-native proteins (such as *B. thuringiensis* crystal proteins) is becoming a critical step in the formulation of plant varieties which express insect resistance genes. In recent years considerable research has yielded tools for the manipulation of endotoxin-encoding genes to permit expression of their encoded proteins *in planta*. Scientists have shown that

10   maintaining a significant level of an mRNA species in a plant is often a critical factor. Unfortunately, the causes for low steady state levels of mRNA encoding foreign proteins are many. First, full-length RNA synthesis may not occur at a high frequency. This could, for example, be caused by the premature termination of RNA during transcription or due to unexpected mRNA processing during transcription. Second, full-length RNA

15   may be produced in the plant cell, but then processed (splicing, polyA addition) in the nucleus in a fashion that creates a nonfunctional mRNA. If the RNA is not properly synthesized, terminated and polyadenylated, it cannot move to the cytoplasm for translation. Similarly, in the cytoplasm, if mRNAs have reduced half lives (which are determined by their primary or secondary sequence) inisufficient protein product will be

20   produced. In addition, there is an effect, whose magnitude is uncertain, of translational efficiency on mRNA half-life. In addition, every RNA molecule folds into a particular structure, or perhaps family of structures, which is determined by its sequence. The particular structure of any RNA might lead to greater or lesser stability in the cytoplasm. Structure *per se* is probably also a determinant of mRNA processing in the nucleus.

25   Unfortunately, it is impossible to predict, and nearly impossible to determine, the structure of any RNA (except for tRNA) *in vitro* or *in vivo*. However, it is likely that dramatically changing the sequence of an RNA will have a large effect on its folded structure It is likely that structure *per se* or particular structural features also have a role in determining RNA stability.

To overcome these limitations in foreign gene expression, researchers have identified particular sequences and signals in RNAs that have the potential for having a specific effect on RNA stability. In certain embodiments of the invention, therefore, there is a desire to optimize expression of the disclosed nucleic acid segments *in planta*. One particular method of doing so, is by alteration of the bacterial gene to remove sequences or motifs which decrease expression in a transformed plant cell. The process of engineering a coding sequence for optimal expression *in planta* is often referred to as "plantizing" a DNA sequence.

Particularly problematic sequences are those which are A+T rich. Unfortunately, since *B. thuringiensis* has an A+T rich genome, native crystal protein gene sequences must often be modified for optimal expression in a plant. The sequence motif ATTTA (or AUUUA as it appears in RNA) has been implicated as a destabilizing sequence in mammalian cell mRNA (Shaw and Kamen, 1986). Many short lived mRNAs have A+T rich 3' untranslated regions, and these regions often have the ATTTA sequence, sometimes present in multiple copies or as multimers (*e.g.*, ATTTATTTA...). Shaw and Kamen showed that the transfer of the 3' end of an unstable mRNA to a stable RNA (globin or VA1) decreased the stable RNA's half life dramatically. They further showed that a pentamer of ATTTA had a profound destabilizing effect on a stable message, and that this signal could exert its effect whether it was located at the 3' end or within the coding sequence. However, the number of ATTTA sequences and/or the sequence context in which they occur also appear to be important in determining whether they function as destabilizing sequences. Shaw and Kamen showed that a trimer of ATTTA had much less effect than a pentamer on mRNA stability and a dimer or a monomer had no effect on stability (Shaw and Kamen, 1987). Note that multimers of ATTTA such as a pentamer automatically create an A+T rich region. This was shown to be a cytoplasmic effect, not nuclear. In other unstable mRNAs, the ATTTA sequence may be present in only a single copy, but it is often contained in an A+T rich region. From the animal cell data collected to date, it appears that ATTTA at least in some contexts is important in stability, but it is not yet possible to predict which occurrences of ATTTA are destabiling elements or whether any of these effects are likely to be seen in plants.